视觉工作记忆离线态表征具有稳固性优势*

李子媛[1] 任国防[1] 袁子昕[2] 喻青青[1] 伍越[1] 刘强[1,3]

（[1]安阳师范学院教育学院，安阳 455000）（[2]西南财经大学工商管理学院，成都 611130）（[3]四川师范大学脑与心理科学研究院，成都 610066）

摘 要 视觉工作记忆状态模型提出了一个层级化的存储结构，将存储态分为在线态和离线态。先前研究已证明，无论是在线态还是离线态的记忆表征都表现出一定程度的稳固性，能够抵制外部因素干扰。然而，在相同的认知加工下，这两种存储状态的记忆表征是否具有同等的稳固性能，目前缺乏明确的认识。为探究这一问题，本研究采用能够同时检测在线态和离线态表征的序列呈现记忆任务范式，并调控干扰刺激呈现（实验 1）和记忆维持时间（实验 2）两个变量，来评估这两种因素分别对不同存储状态的影响。结果显示，在多存储态记忆维持期间呈现干扰刺激或延长记忆维持时间，会显著损害在线态记忆成绩，而不影响离线态成绩。这一结果表明，相比在线态记忆表征，离线态表征在抵制无关干扰和时间衰减方面具有稳固性优势。

关键词 视觉工作记忆，状态模型，在线态，离线态，稳固性优势

分类号 B842

1 前言

视觉工作记忆在认知活动中扮演着至关重要的角色，它不仅负责信息的临时存储，还需在记忆维持过程中保持信息的精确性，以确保高级认知活动得以顺利进行(Buschman & Miller, 2023; Miller et al., 2018; Teng & Kravitz, 2019; Oliver et al., 2011)。然而，记忆信息的稳定维持并非易事。在这一过程中，记忆信息不可避免地遭受损失。无关信息的干扰和信息随时间的自然衰减是导致记忆产生损耗的两个主要因素(Bae & Luck, 2019; Hakim et al., 2021;Liesefeld et al., 2020; Lorenc et al., 2021; Rademaker et al., 2019)。前者可能源自外部环境，例如噪音、干扰物等，这些干扰因素会分散我们的注意力，导致信息的丢失。后者则是指，随着时间的推移，存储在工作记忆中的信息会逐渐变得模糊不清，这一现象可能是由记忆的生理机制或神经生物学因素所导致，因此，信息随时间的自然衰减也是一个不容忽视的损失源。深入理解这些因素对记忆信息维持的影响，对于揭示记忆信息存储的认知机制具有重要意义。这不仅有助于我们更好地理解工作记忆的运行原理，还能提高认知加工的效率和准确

收稿日期：2024-05-18

* 国家自然科学基金项目(31970989)；四川省自然科学基金（2023NSFSC0123）；河南省高等教育教学改革研究与实践项目（2024SJGLX1074）。

通讯作者：刘强, E-mail: lq780614@163.com

性。

工作记忆传统理论模型认为，记忆信息的维持需要依赖持续的神经放电活动。然而，随着研究的不断深入，学者们逐渐认可无神经激活的记忆表征存储模式，即以突触权重变化为基础的离线态记忆。这种认识促进了工作记忆状态理论模型的形成(Cowan, 1999; Kozachkov et al., 2022; Masse et al., 2019; Muhle-Karbe et al., 2021; Oberauer, 2002)，该模型将工作记忆存储表征态分为在线态和离线态，并且认为记忆表征根据任务需求在这两种存储态之间灵活切换。对于多个项目的记忆维持，当前正被关注的项目处于在线态，用于正在进行的认知操作，可通过信号检测设备观测到其诱发的神经活动。当这些项目与当前任务暂时无关，但需要在随后探测中被提取时，他们会进入离线态，通过突触联结来进行记忆信息的短暂维持。记忆表征由在线到离线的存储态切换标定其神经活动降落至基线水平。尽管离线态表征无法通过常规设备被直接探测，但是，以往研究发现在离线态记忆信息的“沉默”存储阶段，向被试呈现强视觉刺激或向离线存储相关脑区施加 TMS 刺激，能观察到无神经活动的离线态表征被重新激活，使其获得可解码性。然而，对于没有神经活动的被遗忘信息(forgotten information)来说，无论施加何种刺激，都观察不到神经信号的重新激活效应(Rose et al., 2016; Wolff et al., 2015; Wolff et al., 2017)。

虽然工作记忆维持和知觉加工在脑区上有所重叠(D'Esposito & Postle, 2015; Kreither et al., 2022; Postle, 2006; Panichello & Buschman, 2021; Serences, 2016; Xu, 2017)，但研究表明，在工作记忆维持期间，呈现与任务无关的视觉刺激并不会显著影响记忆信息的在线维持(Hakim et al., 2020; Kreither et al., 2022; Wang et al., 2018)；同时，与记忆维持相关的神经活动在这一过程中显著增强(Bonnefond & Jensen, 2012; Wang et al., 2022)。这似乎表明，伴随神经激活的在线态表征能够通过增强神经激活形成自我保护机制，使其具有抵制无关干扰的能力。正如以往研究利用回溯线索来调控记忆项目存储神经态，结果发现，那些当前关注的在线态记忆项目能够有效抵制时间衰减和无关信息干扰(Myers et al., 2018; Rerko & Oberauer, 2013; Souza et al., 2016)。需要注意的是，相似的表征稳固性也体现在离线态存储模式上(Wan et al., 2020; Stokes et al., 2020; Zhang et al., 2022)。在 Zhang 等人(2022)关于离线态存储机制的研究中，作者通过记录被试在进行序列记忆任务时的脑电信号，来确定第一个记忆阵列在第二个记忆阵列的在线存储期间以离线态模式进行存储。行为结果发现，在第一个阵列的离线存储阶段，无论延长记忆表征的维持时间，还是呈现探测阵列（知觉干扰）或决策响应（认知干扰），离线态记忆成绩几乎不受影响。这一结果表明，离线态表征能够抵制这些因素的影响。虽然以往研究者采用不同实验任务证明了不同存储模式下记忆表征具有一定的稳固性，但是这些记忆表征的特性一直以来被独立考察，因此无法明晰在相同认知处理下，不同存储状态的记忆表征在稳固性上是否存在差异。对这一问题的探索能够深化我们对工作记忆在线态和离线态表征特性的理解。

因此，本研究将通过实验比较工作记忆在线态和离线态表征的存储稳固性，以确定

哪种存储状态具有显著优势。我们将采用能够有效同时检测这两种存储状态的序列呈现记忆任务范式(Zhang et al., 2022; Li et al., 2020, 2021)，并通过在两种状态的工作记忆表征维持期间调控干扰刺激呈现（实验 1a/1b）和记忆维持时间（实验 2a/2b），确保不同存储状态下的记忆表征经历相同程度的干扰和维持时间，进而考察这两种因素对不同存储状态记忆表征的影响。我们预期，如果这两种存储状态的记忆表征具有同等的稳固性，那么无论呈现干扰刺激还是延长维持时间，这两种存储状态下的记忆成绩都不会显著下降。而如果其中一存储状态记忆表征的稳固性较弱，那么干扰呈现或维持时间延长将导致该存储状态的记忆成绩显著下降。

2　实验 1 ：无关干扰对不同存储状态记忆表征的影响

本实验主要探究无关干扰对不同存储状态记忆表征的影响，为此，在工作记忆多状态维持期间呈现干扰刺激。如果在线态和离线态表征具有同等的稳固性能，那么干扰刺激对这两种存储状态的记忆成绩均无显著影响；而如果其中一存储状态的稳固性较弱，那么干扰刺激的出现将导致该状态的记忆成绩显著下降。

2.1 实验 1a

2.1.1 被试

通过 G*Power 3.1 软件计算所需样本量。根据效应量(dz)为 0.65，α水平为 0.05，为达到 80%的统计检验力所需最小样本量为 28。30 名大学生参加实验 1a，包括 5 名男性，平均年龄 22.03 岁(SD: 2.51)。所有被试的视力或矫正视力正常，实验开始前均签署知情同意书。实验后可获得 30 元作为报酬。所有实验均符合赫尔辛基宣言，并获得四川师范大学人类研究机构审查委员会批准。

2.1.2 实验材料和设备

我们使用 Matlab2014 运行实验程序。电脑大小为 1920×1080 像素，60HZ 刷新率，屏幕背景为灰色(125×125×125)，屏幕中心始终呈现一个黑色注视点(0.23°)。注视点周围出现的正方形色块作为记忆刺激(0.49°×0.49°)，其颜色从蓝(0, 0, 255)、红(255, 0, 0)、橙(255, 165, 0)、品红(255, 0, 255)、绿(0, 255, 0)、青(0, 255, 255)、褐(139, 69, 19)、黄(255, 255, 0)、紫(160, 32, 240)、靛蓝(75, 0, 130)这十种颜色中随机选择。在同一个试次中，不重复使用同一种颜色作为记忆刺激。记忆阵列 1 和 2 中的 4 个色块分布在一个以注视点为中心，半径为 2.02°的隐形圆周上。每个阵列中的记忆色块之间等距，空间位置不重叠（如图 1）。距离注视点 0.61°的虚拟圆上等距呈现 6 个圆形色块（0.28°），作为干扰刺激。其颜色 RGB 为(216, 191, 216), (100, 149, 237), (128, 128, 0), (210, 105, 30), (160, 82, 45), (255, 255, 255)。每种颜色随机出现在 6 个位置上。两个探测阵列依次呈现，分别检测记忆阵列 2 和 1。探测 1 和 2 中的黑色边框方块的位置和数量分别与记忆阵列 2 和 1 相对应，其中涂上颜色的色块作为探测刺激。视距固定为 70 厘米。

2.1.3 实验设计

如图 1a，记忆阵列 1 包含 4 个色块，呈现时间为 0.5s。1s 的空白间隔之后，记忆阵列 2 出现，也包含 4 个色块，呈现时间为 0.5s。在干扰条件下，1s 间隔之后，干扰刺激出现，其持续时间为 0.2s。为避免干扰刺激在知觉层面对探测 1 产生影响，0.8s 的延迟间隔之后，探测阵列 1 出现。在无干扰条件下，记忆阵列 2 消失之后，空白屏幕持续 2s 后探测阵列出现。由于预期内的干扰物呈现会使被试有意识地增强对记忆项目的保护(Wang et al., 2022)，因此，干扰刺激以 30%的概率在记忆维持期间呈现，并要求被试对其忽略。被试需要判断探测 1 中出现的色块颜色是否与记忆阵列 2 相应位置上的色块颜色相同，如果相同，则按“F”键，否则按“J”键。探测 1 消失后经过 0.8s 的延迟间隔，探测 2 出现，要求被试判断该探测阵列中的色块颜色是否和记忆阵列 1 中相应位置上的颜色相同。两个探测阵列只有在被试做出按键反应之后，才会消失。因此，在进行记忆任务过程中，被试应该注重正确率而非速度。每个组块中探测颜色发生变化或不变化的概率相等。为排除特定记忆项目空间位置对行为结果的影响，两个记忆阵列中项目的空间位置并非始终固定。当记忆阵列 1 的空间构型为“十字”型时（如图 1a 记忆阵列 1），记忆阵列 2 的空间构型为“正方”型（如图 1a 记忆阵列 2），该呈现方式为类型 1；当记忆阵列 1 的空间构型为“正方”型时，记忆阵列 2 的空间构型为“十字”型，该呈现方式为类型 2。一半被试完成类型 1 的呈现方式，另一半被试完成类型 2 呈现方式。

每个被试需要完成 8 个组块，一共 256 个试次。组块之间至少有 30s 的休息时间。在正式实验开始前，要求被试先阅读指导语，并完成至少 12 个练习试次，确保他们熟悉整个实验流程，同时要求禁止采用语音编码策略。

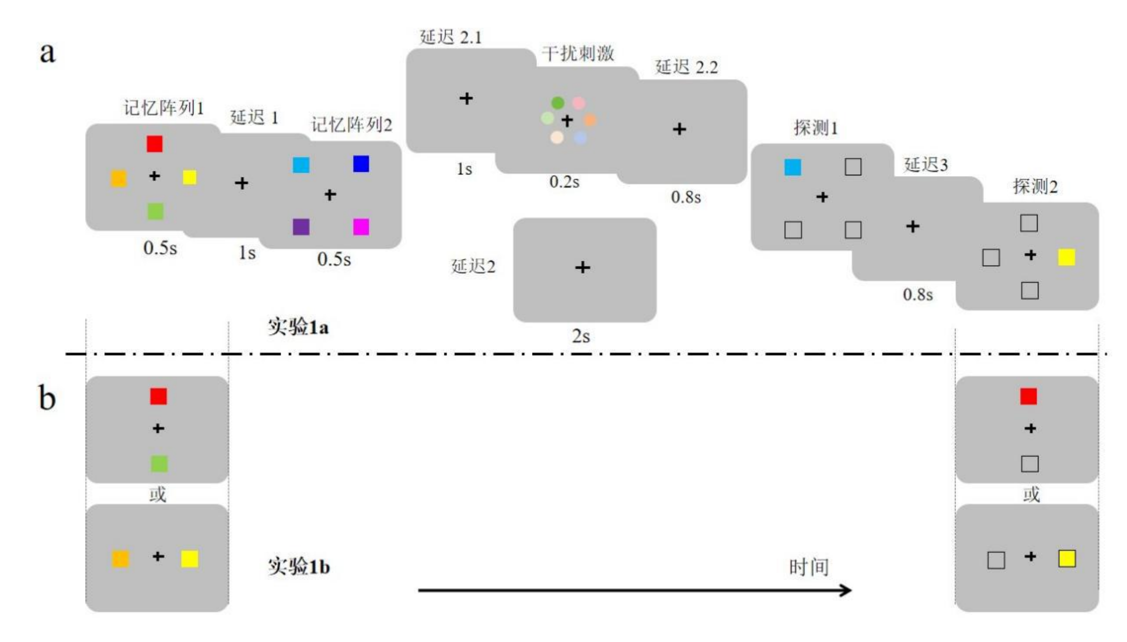


图 1　实验 1 流程示意图

2.1.4 数据分析和结果

我们主要关注两个记忆阵列的正确率。探测 1 和 2 分别反映记忆阵列 2（在线态记忆）和记忆阵列 1（离线态记忆）成绩。所有被试的行为结果进行 2（记忆类型：在线态 vs. 离线态）×2（干扰条件：有干扰 vs. 无干扰）重复测量方差分析。如果交互作用显著，再分别对不同存储状态的记忆成绩进行配对样本 t 检验。结果如图 2a 所示。两个阵列的记忆成绩均高于随机水平，表明所有记忆项目被有效存储。重复测量方差分析结果显示，记忆类型(F (1, 29) = 19.33, p < 0.001, $\eta_p{}^2$ = 0.40)和干扰条件(F (1, 29) = 17.90, p = 0.001, $\eta_p{}^2$ = 0.38)的主效应均显著。重要的是，这两个因素之间存在显著的交互作用，F (1, 29) = 8.98, p = 0.006, $\eta_p{}^2$ = 0.24。在随后的配对样本 t 检验结果中发现，相比无干扰条件，在线态记忆成绩在有干扰条件下显著下降，t (29) = 4.38, p < 0.001, Cohen’s d = 0.80, 95% CI [-1.21, -0.38]；而离线态记忆成绩在有、无干扰条件之间无显著差异，t (29) = 0.89, p = 0.38, Cohen’s d = 0.16, 95% CI [-0.52, 0.20]。

以上结果表明在工作记忆多状态维持期间，随机呈现的干扰刺激会显著损害在线态记忆成绩，而不影响离线态记忆的维持。据此可推断，在抵制无关干扰方面，离线态记忆表征具有稳固性优势。

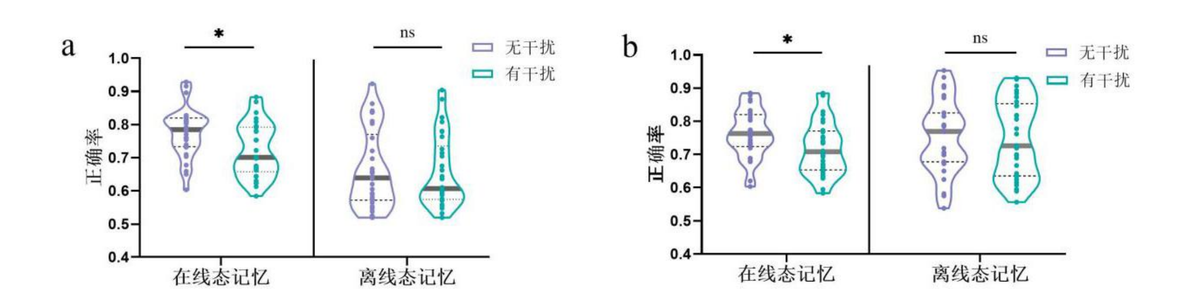


图 2　实验 1a（a）和实验 1b（b）的结果图

注：*表示 p < 0.05, ns 表示 p > 0.05

2.2　实验 1b

以往研究发现，在工作记忆多状态存储过程中，记忆项目从在线态转移到离线态会产生信息衰减(Zhang et al., 2022)。因此，我们在实验 1a 中观察到相比在线态记忆，离线态记忆（记忆阵列 1）表现相对较差。这就引发了一种可能性，即已经衰减到较低水平的离线态记忆可能对无关干扰不敏感。为排除这一可能性，我们通过降低离线态负荷来提高其记忆表现，使其与在线态记忆水平相当。在这种情况下，如果离线态记忆成绩在有、无干扰条件下仍然无显著差异，那么可进一步确定离线态表征在抵制无关干扰方面的稳固性优势。

2.2.1 研究方法

实验 1b 招募 29 名大学生参加该实验，包括 7 名男性，平均年龄 20.55 岁(SD:1.40)。实验设备，环境、刺激和实验流程参照实验 1a，不同的是记忆阵列 1 只包含 2 个项目（见图 1b）。这些项目以注视点为中心垂直或水平呈现。记忆阵列 2 包含 4 个项目，其空间构型为

“正方”型。

2.2.2 数据分析和结果

数据分析方法同实验 1a，结果如图 2b 所示。2（记忆类型：在线态 vs. 离线态）×2（干扰条件：有干扰 vs. 无干扰）重复测量方差分析结果显示，干扰条件的主效应显著，$F(1, 28) = 10.06$, $p = 0.004$, $\eta_p^2 = 0.26$，记忆类型的主效应不显著，$F(1, 28) = 0.29$, $p = 0.59$, $\eta_p^2 = 0.01$。记忆类型和干扰条件之间有显著的交互作用，$F(1, 28) = 8.12$, $p = 0.008$, $\eta_p^2 = 0.23$。随后进行配对样本 t 检验，结果显示相比无干扰条件，在线态记忆成绩在有干扰条件下显著下降，$t(28) = 4.32$, $p < 0.001$, Cohen’s d = 0.80, 95% CI [-1.22, -0.38]；而离线态记忆成绩在有、无干扰条件之间无显著差异，$t(28) = 0.53$, $p = 0.60$, Cohen’s d = 0.10, 95% CI [-0.46, 0.27]。这些结果表明，当提高离线态记忆成绩使其与在线态记忆水平相当时，随机呈现的干扰刺激仍然不影响离线态记忆表征的维持，但是显著损害在线态记忆。该结果进一步验证了实验 1a 的结论，即相比在线态，离线态表征在抵制无关干扰方面具有稳固性优势。

2.3 讨论

实验 1a/1b 结果一致表明在抵制无关干扰方面，离线态记忆表征比在线态具有更强的稳固性能。需要注意的是，在当前采用的序列呈现记忆任务中，干扰刺激出现在两个记忆阵列呈现后的延迟期间，那么干扰刺激在时间上更接近阵列 2（在线态记忆），而相对远离阵列 1（离线态记忆）。这一时间效应可能导致在线态记忆更易受干扰刺激的影响。然而， Zhang 等人(2022)的实验结果发现，在阵列 2 呈现之前，阵列 1 的记忆维持诱发持续的记忆相关脑电成分 CDA。这表明在阵列 2 出现之前，阵列 1 的信息一直维持在在线态。只有在阵列 2 呈现后，阵列 1 的表征才被转换到离线态。那么阵列 1 的离线态表征和阵列 2 的在线态表征出现的时间应该是近似同步的。因此，在干扰条件下，在线态成绩的下降并非由干扰刺激的时间效应导致，而是由于在线态表征的稳固性相对较弱。

3 实验 2：时间衰减对不同存储态记忆表征的影响

实验 2 主要考察时间衰减对不同存储状态记忆表征的影响。为此，该实验调控记忆表征多状态维持期间的记忆时间（1s vs. 3s）。我们预测，如果两个存储状态的记忆表征在抵制时间衰减方面具有同等的稳固性，那么延长记忆维持时间不会使这两个存储状态的记忆成绩产生衰减。相反，如果其中一个存储状态在时间衰减方面稳固性较弱，那么该状态的记忆成绩随维持时间延长而显著下降。

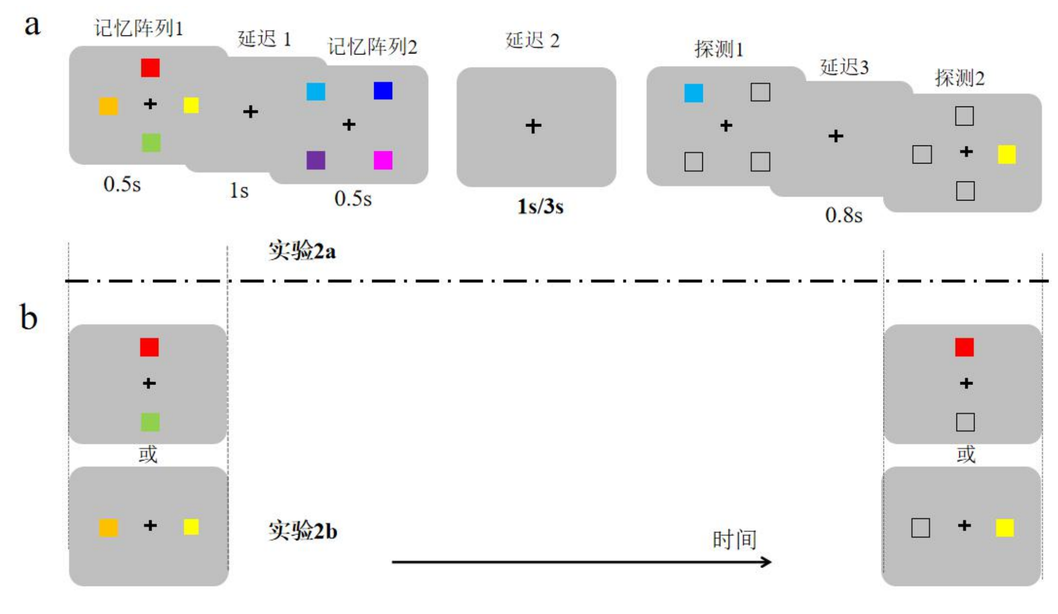


图 3　实验 2 流程示意图

3.1 实验 2a

3.1.1 研究方法

本实验招募 28 名被试，其中包括 11 男性，平均年龄 21.32 岁(SD：1.79)。实验设备，环境和刺激与实验 1a 保持一致。在此基础上，我们作了如下修改：如图 3a 所示，记忆阵列 2 之后的延迟 2 是主要调控变量，分别为 1s 和 3s。该实验为被试内设计，每个被试需要完成长（3s）、短（1s）维持时间条件，顺序在被试间平衡。每种条件包含 3 个组块，一共 192 个试次。组块之间至少有 30s 的休息时间。在正式实验开始前，要求被试先阅读指导语，并完成至少 12 个练习试次，确保他们熟悉整个实验流程，同时要求他们禁止采用语音编码策略。

3.1.2 数据分析和结果

记忆行为正确率是我们主要关注的测量指标。所有被试的行为结果进行 2（记忆类型：在线态 vs. 离线态）× 2（维持时间：1s vs. 3s）重复测量方差分析。如果交互作用显著，再进行简单效应分析，结果如图 4a 所示。两个记忆阵列成绩均高于随机水平。重复测量方差分析结果显示，记忆类型(F (1,27) = 37.40, p < 0.001, η_p^2 s = 0.58)和维持时间(F (1,27) = 8.72, p = 0.001, η_p^2 = 0.58)的主效应均显著。重要的是，记忆类型和维持时间之间的交互作用也显著，F (1, 27) = 5.68, p = 0.02, η_p^2 = 0.17。随后的配对样本 t 检验结果显示，对于在线态记忆成绩，间隔时间从 1 秒延长至 3 秒，正确率显著下降，t (27) = 3.17, p = 0.004, Cohen’s d = 0.60, 95% CI [0.19, 0.10]。但是，长、短间隔条件下的离线态成绩无显著差异，t (27) = 0.86, p = 0.40, Cohen’s d = 0.16, 95% CI [-0.21, 0.54]。这些结果表明，在抵制时间衰减方面，在线态表征表现出一定的易感性，而离线态表征具有稳固性优势。

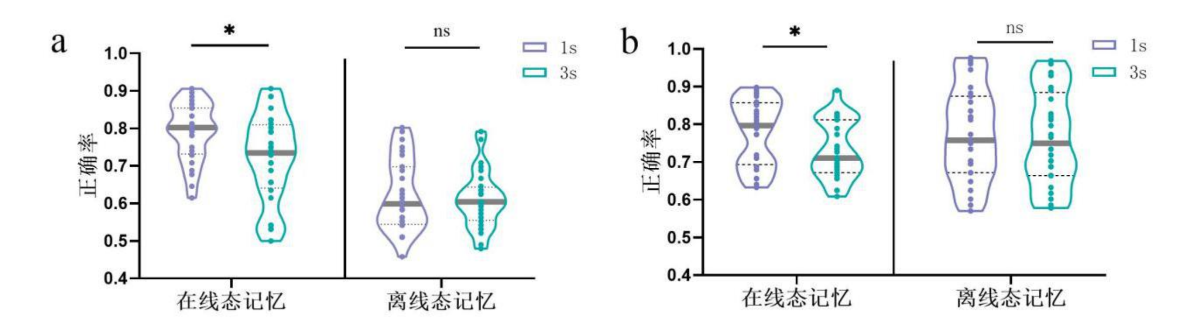


图 4　实验 2a（a）和实验 2b（b）的结果图

注：*表示 p < 0.05, ns 表示 p > 0.05

3.2　实验 2b

在实验 2a 中，由于状态切换而导致的记忆衰减使得离线态记忆表现相对较差。在这种情况下，记忆阵列 1 对时间衰减不敏感同样可能是由于其成绩已经衰减至较低水平。为排除这种可能性，采取和实验 1b 一样的调控方式，将记忆阵列 1 的负荷降低至 2 个项目来提高其表现水平。如果相比在线态记忆表征，离线态表征在抵制时间衰减方面确实具有稳固性优势，那么维持时间无论是 1 秒还是 3 秒，提高记忆表现水平的离线态记忆仍然无显著变化，而在线态记忆成绩随维持时间延长而下降。

3.2.1 研究方法

28 名大学生参加该实验，包括 5 名男性，平均年龄 20.68 岁(SD:2.19)。实验设备，环境、刺激和实验流程参照实验 2a，不同的是记忆阵列 1 只包含 2 个项目（如图 3b）。这些项目以注视点为中心垂直或水平呈现。记忆阵列 2 包含 4 个项目，其空间构型为“正方”型。

3.2.2 数据分析和结果

数据分析方法同实验 2a，结果如图 4b 所示。2（记忆类型：在线态 vs. 离线态）× 2（维持时间：1s vs. 3s）重复测量方差分析结果显示，记忆类型的主效应不显著，F (1, 27) = 0.27, p = 0.61, $\eta_p{}^2$ = 0.01。维持时间的主效应显著，F (1, 27) = 8.16, p = 0.008, $\eta_p{}^2$ = 0.23，并且维持时间和记忆类型之间的交互作用也显著，F (1, 27) = 7.56, p = 0.01, $\eta_p{}^2$ = 0.22。配对样本 t 检验结果显示，对于在线态记忆成绩，间隔时间从 1 秒延长至 3 秒，正确率显著下降，t (27) = 3.92, p < 0.001, Cohen’s d = 0.74, 95% CI [0.32, 1.16]。但是，长、短间隔条件下的离线态记忆成绩无显著差异，t (27) = 0.17, p = 0.87, Cohen’s d = 0.03, 95% CI [-0.34, 0.40]。这些结果表明，当通过降低离线态负荷来提高该存储状态记忆成绩使之与在线态记忆水平相当时，离线态表征在抵制时间衰减方面仍然表现出稳固性优势。

4 总讨论

当前研究旨在比较工作记忆在线态和离线态表征在相同的认知处理下的稳固性。为此，我们通过两个实验来确定哪一存储状态的记忆表征在抵制无关干扰和时间衰减方面具有稳固性优势。实验 1 通过调控多状态存储期间的干扰刺激呈现来考察其对不同存储状态下记忆

表征的影响。结果显示，在干扰呈现条件下，在线态记忆成绩显著降低，而离线态记忆成绩则未受显著影响。实验 2 调控多状态维持期间的记忆维持时间，来检验时间衰减对不同存储状态记忆表征的影响。结果显示，当记忆维持时间从 1 秒延长至 3 秒时，在线态记忆成绩随时间延长而显著降低，而离线态记忆成绩无显著变化。综上，当前研究结果表明，不同存储状态的记忆表征在相同的认知处理下存在稳固性差异。具体来说，在抵制无关干扰和时间衰减方面，离线态记忆表征具有稳固性优势。

需要注意的是，先前研究也观察到离线态表征的稳固性优势(Chota & Van der Stigchel, 2021; Mongillo et al., 2008; Olivers & Roelfsema, 2020)。例如，Hu 等人(2014，2016)采用一系列快速序列呈现任务，并且利用近因效应和奖励机制，优先化特定项目使其处于在线态。结果发现，相比序列中未被优先化的离线态项目，在线态项目被干扰刺激显著影响。然而，在这一序列呈现记忆任务中，虽然表现有近因效应的记忆表征被认为存储在在线态，但它们可能同时受到先前呈现记忆项目的前摄干扰，这使得难以评估干扰刺激对在线态记忆损耗的影响程度。另外，Mallett 和 Lewis-Peacock(2019)采用 fMRI 和多体素模式分析技术，对不同存储状态下记忆表征的特征进行深入探讨。结果显示，在分心物出现后的维持阶段，无论在线态还是离线态记忆表征均能被成功解码。但是，只有离线态表征的解码准确性达到分心物出现前的水平。这表明，离线态表征在稳固性方面比在线态记忆表征更有优势。然而，该研究设计存在一些关键性限制。首先，干扰任务出现之前有 8 秒的延迟期，这可能导致序列呈现的所有记忆项目在干扰任务出现前都已进入离线态(Wang et al., 2018)。其次，研究中未设置无干扰物的对照条件，因此无法确定在线态表征在行为层面上是否同样遭受记忆损耗。综上，尽管先前研究指出了离线态表征可能在稳固性方面具有优势，但这些研究并未提供充分证据来验证这一假设。此外，对于时间衰减对记忆表征影响的探讨相对欠缺。因此，本研究基于工作记忆多状态存储模型，从行为层面系统对比不同存储状态下记忆表征在相同认知处理下的稳固性。本研究不仅弥补了以往研究在理解不同存储状态记忆表征稳固性差异方面的不足，而且为深入探讨工作记忆的存储机制提供了新视角。这些结果强调了离线态在工作记忆维持过程中作为一种有效的记忆保护机制，有助于信息的稳定维持。

4.1 离线态表征稳固性优势的神经机制

离线态表征在抵制无关干扰和时间衰减方面的稳固性优势可能源于其基于突触权重变化的神经机制。离线态表征涉及的突触模式与细胞层面的活动变化有着密切联系，而这些变化并不依赖于持续的神经放电活动(Masse et al., 2019; Mongillo et al., 2008; Stokes, 2015)。这种静态的存储机制能够有效规避与无关干扰刺激相关的神经活动相互作用，从而保护记忆表征不受干扰。此外，离线态表征的突触模式显示了特定神经通路中神经元突触连接的强化，这种状态的维持需要较低的神经能量消耗。与此相对，需要持续神经激活的在线态表征模式则消耗较多的神经能量，可能导致认知资源的枯竭和神经回路的疲劳。因此，离线态表征在时间维度上相较于在线态表征显示出更高的效率和稳定性(Chota & Van der Stigchel, 2021;

Olivers & Roelfsema, 2020; Wang et al., 2022)。当前的研究发现揭示了记忆存储机制的复杂性和多样性，其中在线态表征通常指的是记忆内容在大脑中持续被激活，以维持其当前可用性；而离线态表征则指的是记忆内容在特定时期暂时休眠，以减少能量消耗和防止信息干扰。这两种状态在记忆保持过程中交替出现，共同维护着记忆的稳定性。

4.2 在线态表征的易感性特征

在线态表征通常被认为处于注意焦点中，并通过占用大量认知资源来抵御外部干扰。例如，通过使用回溯线索将注意资源集中到特定的单一项目上，可以显著增强其该项目的稳固性，消除干扰效应(Makovski & Jiang, 2007; Landman et al., 2003; Griffin & Nobre, 2003)。但是正如当前结果显示，这些表征有时也表现出一定的易感性(Souza et al., 2014; Souza & Oberauer, 2016; Allen & Ueno, 2018)。基于这些发现，我们推测在线态记忆表征的稳固性和易感性（protective vs. susceptible）受到多种因素影响，包括项目数量，线索类型以及被试对线索的策略性使用(van Moorselaar et al., 2015; Heuer & Schubö, 2016; Delvenne & Holt, 2012)。当在线态同时维持多个记忆项目时，这些表征可能表现出较弱的稳固性，而干扰刺激的出现导致记忆损耗。相反，当只有一个项目以在线态模式被维持时，该表征的稳固性则显著增强。这表明记忆表征的稳固性并非固定不变，而是受记忆任务具体情境的影响。在当前研究的多状态存储阶段，在线态同时存储 4 个记忆项目。这种高负荷条件下，维持阶段的干扰因素可能导致在线态记忆信息的损耗。那么当在线态记忆负荷减少至单个项目时，干扰刺激是否仍然会对在线态表征产生损害效应。这一问题值得进一步探讨。

4.3 记忆表征稳固性的潜在影响因素

工作记忆在信息维持阶段对干扰刺激的预期性能够显著调控记忆信息的表征强度。Wang 等人(2022)通过操纵干扰刺激出现的概率，有效地调控了干扰的预期性。研究发现，当记忆项目的特征信息预期将遭受干扰时，该信息的表征强度得到增强；而那些预期不会受到干扰的特征信息则相对脆弱，更易受到损害。功能性磁共振成像(fMRI)数据进一步揭示，在视觉皮层中，预期遭受干扰的特征信息的分类器解码正确率高于随机水平，表明预期干扰增强了这些信息的神经表征。这些发现指出，干扰的可预期性可能促使被试无意识地增强特定记忆信息的表征强度，作为一种保护性机制，以提升其稳固性。为了排除干扰预期对记忆表征强度的潜在增强作用，本研究以小概率方式呈现干扰刺激，旨在减少被试通过干扰预期而增强记忆表征的可能性。然而，这种做法引发了新的研究问题：如果干扰刺激的呈现具有高度可预期性，被试是否会有意识地增强在线态表征的强度，从而提升其稳固性？此外，干扰刺激的可预期性是否也可能影响以突触模式存储的离线态表征的强度？这些问题为未来研究提供了新的探索方向，需要进一步的实验来阐明。

5 结论

综上，本研究采用能够同时检测工作记忆在线态和离线态记忆表征的序列呈现记忆任务范式，对比不同存储状态下记忆表征在相同认知处理下的稳固性。研究发现，相比在线态，

离线态表征在抵制无关干扰和时间衰减方面具有稳固性优势。这一发现不仅为理解工作记忆的存储机制提供了新的视角，而且丰富了我们对工作记忆多状态存储机制的认识，为未来的记忆研究提供了新的方向和思路。

参考文献

Allen, R. J., & Ueno, T. (2018). Multiple high-reward items can be prioritized in working memory but with greater vulnerability to interference. *Attention Perception & Psychophysics, 80*(7), 1731−1743.

Bae, G. Y., & Luck, S. J. (2019). What happens to an individual visual working memory representation when it is interrupted? *British Journal of Psychology, 110*(2), 268−287.

Bonnefond, M., & Jensen, O. (2012). Alpha oscillations serve to protect working memory maintenance against anticipated distracters. *Current Biology,* 22*(20)*, 1969−1974.

Buschman, T. J., & Miller, E. K. (2023). Working memory is complex and dynamic, like your thoughts. *Journal of Cognitive Neuroscience, 35*(1), 17−23.

Chota, S., & Van der Stigchel, S. (2021). Dynamic and flexible transformation and reallocation of visual working memory representations. *Visual Cognition, 29*(7), 409−415.

Cowan, N. (1999). An embedded-processes model of working memory. *Models of working memory: Mechanisms of active maintenance and executive control, 20*(506), 1013−1019.

Delvenne, J. F., & Holt, J. L. (2012). Splitting attention across the two visual fields in visual short-term memory. *Cognition, 122*(2), 258−263.

D'Esposito, M., & Postle, B. R. (2015). The cognitive neuroscience of working memory. *Annual Review of Psychology, 66*(1), 115−142.

Griffin, I. C., & Nobre, A. C. (2003). Orienting attention to locations in internal representations. *Journal of Cognitive Neuroscience, 15*(8), 1176−1194.

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2020). Perturbing neural representations of working memory with task-irrelevant interruption. *Journal of Cognitive Neuroscience, 32*(3), 558−569.

Hakim, N., Feldmann-Wüstefeld, T., Awh, E., & Vogel, E. K. (2021). Controlling the flow of distracting information in working memory. *Cerebral Cortex, 31*(7), 3323−3337.

Heuer, A., & Schubö, A. (2016). Feature-based and spatial attentional selection in visual working memory. *Memory & Cognition, 44*, 621−632.

Hu, Y., Allen, R. J., Baddeley, A. D., & Hitch, G. J. (2016). Executive control of stimulus-driven and goal-directed attention in visual working memory. *Attention, Perception, & Psychophysics, 78*(7), 2164−2175.

Hu, Y., Hitch, G. J., Baddeley, A. D., Zhang, M., & Allen, R. J. (2014). Executive and perceptual attention play different roles in visual working memory: evidence from suffix and strategy effects. *Journal of Experimental Psychology: Human Perception and Performance, 40*(4), 1665−1678.

Kozachkov, L., Tauber, J., Lundqvist, M., Brincat, S. L., Slotine, J. J., & Miller, E. K. (2022). Robust and brain-like working memory through short-term synaptic plasticity. *PLoS Computational Biology, 18*(12), e1010776.

Kreither, J., Papaioannou, O., & Luck, S. J. (2022). Active working memory and simple cognitive operations. *Journal of Cognitive Neuroscience, 34*(2), 313−331.

Landman, R., Spekreijse, H., & Lamme, V. A. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research, 43*(2), 149−164.

Li, Z., Liang, T., & Liu, Q. (2021). The storage resources of the active and passive states are independent in visual working memory. *Cognition, 217*, 104911.

Li, Z., Zhang, J., Liang, T., Ye, C., & Liu, Q. (2020). Interval between two sequential arrays determines their storage state in visual working memory. *Science Reports, 10*(1), 7706.

Liesefeld, H. R., Liesefeld, A. M., Sauseng, P., Jacob, S. N., & Müller, H. J. (2020). How visual working memory handles distraction: cognitive mechanisms and electrophysiological correlates. *Visual Cognition, 28*(5−8), 372−387.

Lorenc, E. S., Mallett, R., & Lewis-Peacock, J. A. (2021). Distraction in visual working memory: resistance is not futile. *Trends in Cognitive Sciences, 25*(3), 228−239.

Makovski, T., & Jiang, Y. V. (2007). Distributing versus focusing attention in visual short-term memory. *Psychonomic Bulletin & Review, 14*(6), 1072−1078.

Mallett, R., & Lewis-Peacock, J. A. (2019). Working memory prioritization impacts neural recovery from distraction. *Cortex, 121*, 225−238.

Masse, N. Y., Yang, G. R., Song, H. F., Wang, X. J., & Freedman, D. J. (2019). Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nature Neuroscience, 22*(7), 1159−1167.

Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working memory 2.0. *Neuron, 100*(2), 463−475.

Mongillo, G., Barak, O., & Tsodyks, M. (2008). Synaptic theory of working memory. *Science, 319*(5869), 1543−1546.

Muhle-Karbe, P. S., Myers, N. E., & Stokes, M. G. (2021). A hierarchy of functional states in working memory. *Journal of Neuroscience, 41*(20), 4461−4475.

Myers, N. E., Chekroud, S. R., Stokes, M. G., & Nobre, A. C. (2018). Benefits of flexible prioritization in working memory can arise without costs. *Journal of Experimental Psychology: Human Perception and Performance, 44*(3), 398−411.

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 411−421.

Olivers, C. N., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: when it guides attention and when it does not. *Trends in Cognitive Science, 15*(7), 327−334.

Olivers, C. N. L., & Roelfsema, P. R. (2020). Attention for action in visual working memory. *Cortex, 131*, 179−194.

Panichello, M. F., & Buschman, T. J. (2021). Shared mechanisms underlie the control of working memory and attention. *Nature, 592*(7855), 601−605.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience, 139*(1), 23−38.

Rademaker, R. L., Chunharas, C., & Serences, J. T. (2019). Coexisting representations of sensory and mnemonic information in human visual cortex. *Nature Neuroscience, 22*(8), 1336−1344.

Rerko, L., & Oberauer, K. (2013). Focused, unfocused, and defocused information in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(4), 1075−1096.

Rose, N. S., LaRocque, J. J., Riggall, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science, 354*(6316), 1136−1139.

Serences, J. T. (2016). Neural mechanisms of information storage in visual short-term memory. *Vision Research, 128*, 53−67.

Souza, A. S., & Oberauer, K. (2016). In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Attention, Perception, & Psychophysics, 78*(7), 1839−1860.

Souza, A. S., Rerko, L., & Oberauer, K. (2014). Unloading and reloading working memory: attending to one item frees capacity. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3),

1237-1256.
Souza, A. S., Rerko, L., & Oberauer, K. (2016). Getting more from visual working memory:Retro-cues enhance retrieval and protect from visual interference. *Journal of Experimental Psychology: Human Perception and Performance, 42*(6), 890−910.
Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends in Cognitive Science, 19*(7), 394−405.
Stokes, M. G., Muhle-Karbe, P. S., & Myers, N. E. (2020). Theoretical distinction between functional states in working memory and their corresponding neural states. *Visual Cognition, 28*(5−8), 420−432.
Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour, 3*(8), 827−836.
van Moorselaar, D., Gunseli, E., Theeuwes, J., & NL Olivers, C. (2015). The time course of protecting a visual memory representation from perceptual interference. *Frontiers in Human Neuroscience, 8*, 1053.
Wan, Q., Cai, Y., Samaha, J., & Postle, B. R. (2020). Tracking stimulus representation across a 2-back visual working memory task. *Royal Society Open Science, 7*(8), 190228.
Wang, B., Knapen, T., & Olivers, C. N. L. (2022). Visual working memory adapts to the nature of anticipated interference. *Journal of Cognitive Neuroscience, 34*(7), 1148−1163.
Wang, B., Theeuwes, J., & Olivers, C. N. L. (2018). When shorter delays lead to worse memories: Task disruption makes visual working memory temporarily vulnerable to test interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(5), 722−733.
Wolff, M. J., Ding, J., Myers, N. E., & Stokes, M. G. (2015). Revealing hidden states in visual working memory using electroencephalography. *Frontiers in System Neuroscience, 9*, 123.
Wolff, M. J., Jochim, J., Akyurek, E. G., & Stokes, M. G. (2017). Dynamic hidden states underlying working-memory-guided behavior. *Nature Neuroscience, 20*(6), 864−871.
Xu, Y. D. (2017). Reevaluating the sensory account of visual working memory storage. *Trends in Cognitive Sciences, 21*(10), 794−815.
Zhang, J., Ye, C., Sun, H.-J., Zhou, J., Liang, T., Li, Y., & Liu, Q. (2022). The passive state: a protective mechanism for information in working memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(9), 1235−1248.

Offline memory representations have robustness superiority in visual working memory

LI Ziyuan[1], REN Guofang[1], YUAN Zixin[2], YU Qingqing[1], WU Yue [1], LIU Qiang[1,3]

([1] School of Education, Anyang Normal University, Anyang 455000, China)

([2] Faculty of Business Administeaction, Southwestern University of Finance and Economics, Chengdu 611130, China)

([3] Institute of Brain and Psychological Science, Sichuan Normal University, Chengdu 610066, China)

Abstract

Visual working memory (VWM) plays a crucial role in various higher cognitive activities such as problem-solving, learning, and decision-making (Buschman & Miller, 2023; Teng & Kravitz, 2019; Oliver et al., 2011). The function of VWM extends beyond mere information storage; it also requires the preservation of memory fidelity to ensure the smooth execution of

various cognitive activities (Bae & Luck, 2019; Miller et al., 2018; Liesefeld et al., 2020). State-based models of visual working memory propose a hierarchical structure of memory storage, consisting of both online and offline states. Previous studies have evidenced that memory representations in these two distinct storage states exhibit some resilience to irrelevant distractor and time decay. However, while the properties of memory representations in these two states have been explored independently, it was still unclear whether online and offline memory representations possess comparable resistance to distractor and time decay when exposing to identical cognitive processing.

Thus, this study aimed to compare the robustness of memory representations in the online and offline states, with the goal of determining which storage state had superior resilience. To achieve this, we adopted a sequential presentation memory task, which facilitated the co-existence of both online and offline representations by effectively separating representational storage states. The primary experimental variables were the presence of distractors (Experiment 1) and retention duration (Experiment 2) during memory maintenance involving both states. This design allowed us to examine the effects of these factors on memory representations in the two storage states. A total sample of 115 was recruited in four behavioral experiments. Experiments 1a/1b assessed the effect of distractor on memory representations in the online and offline states, while Experiments 2a/2b examined the effect of time decay. We predicted that, if memory representations in the two states exhibited equal robustness, memory performance in both storage states would remain largely unaffected by the presence of distractors or extended retention duration. Whereas if the representational robustness was weaker in one storage state, then either the presentation of distractor or prolonged retention would result in a significant memory decline in that state.

The results of Experiment 1 collectively showed that when distractor was unexpectedly presented, online memory performance was significantly impaired, while no impairment occurred to offline memory. Experiment 2 jointly revealed that as the retention duration increased from 1s to 3s, significant impairment occurred to online memory only, while offline memory performance was clearly unaffected. These findings indicated that memory representations in the online and offline states differ in their robustness within the identical cognitive processing. Specifically, offline memory had superior resilience compared to the online when it came to resisting both unrelated distractor and time decay.

In summary, during working memory maintenance employing both online and offline storage states, offline memory exhibits a robustness superiority in resisting interference from distractors and temporal decay compared to online memory. These findings contribute to a deeper understanding of multi-state storage mechanisms in working memory, and suggest new directions for future memory research.
Keywords: visual working memory, state-based models, online memory, offline memory, robustness superiority